On the cosmological bias of the excess variance as a variability estimator

Marco Antonucci, Fausto Vagnetti (Università di Roma Tor Vergata), and Dario Trevese (Università di Roma La Sapienza)

1. introduction

The variability of active galactic nuclei (AGN) in the X-ray domain is often estimated with the use of the normalized excess variance (NXS, e.g. Nandra et al. 1997, Vaughan et al. 2003, Ponti et al. 2012, Lanzuisi et al. 2014). This is defined as $\sigma^2_{NXS} = (S^2 - \sigma_n^2)/\overline{x}^2$, where S^2 is the total variance of the light curve, σ_n^2 is the mean square error, and $\overline{x}$ is the mean of N total measurements. This estimator provides a straightforward average variability measure for each source on the basis of the available observations. However, its use is not always appropriate, because this parameter depends on the length of the monitoring time interval Δt_{rest} in the rest-frame of the source, and is therefore biased for cosmological time dilation, as has been pointed out by more authors (e.g. Lawrence & Papadakis 1993, Papadakis et al. 2008, Vagnetti et al. 2011). The bias can be negligible when applied to low-redshift AGNs, provided that uniform monitoring time intervals Δt_{obs} are chosen (see e.g. Ponti et al. 2012). On the contrary, some authors use NXS even for AGNs in wide redshift intervals, so finding uncorrected variability estimates. Ensemble variability studies based on the structure function (SF) have shown that variability in various electromagnetic bands increases with the rest-frame time lag τ_{rest} between observations (e.g. Trevese et al. 1994, deVries et al. 2005, Vagnetti et al. 2011). The SF is a better estimator of the variability, allowing to describe it as a function of the time scale. Defining it as

$$SF(\tau) = \sqrt{\langle[\log f(t+\tau) - \log f(t)]^2\rangle - \sigma_n^2} = \sqrt{\langle(\delta \log f)^2\rangle - \sigma_n^2} \qquad (1)$$

we can estimate the expected value of the normalized excess variance (neglecting the error) in a given time interval $(0, \Delta t_{obs})$ as

$$\sigma^2_{NXS} = [\sigma_f^2/\overline{f}^2]_{(0,\Delta t_{obs})} \approx (\log e)^{-2}\langle\delta \log f^2\rangle_{(0,\Delta t_{obs})}$$

This can be written through the structure function as $(\log e)^{-2}\frac{1}{2}\langle SF^2\rangle_{(0,\Delta t_{rest})}$, where the factor 1/2 accounts for the two independent measurements contributing to each SF flux difference, and $\Delta t_{rest} = \Delta t_{obs}/(1+z)$. Adopting a power-law form $SF = k\tau^{\beta}$ (Vagnetti et al. 2011), the square average in the rest-frame time interval is given by

$$\langle SF^2\rangle = \frac{1}{\Delta t_{rest}}\int_0^{\Delta t_{rest}} k^2\tau^{2\beta}d\tau = \frac{[SF(\Delta t_{rest})]^2}{2\beta+1}$$

so that $\sigma^2_{NXS} = k^2\Delta t_{rest}^{2\beta}/[2(2\beta+1)(\log e)^2]$. It is thus useful to estimate the excess variance for a fixed interval Δt^* as follows:

$$\sigma^2_{NXS}{}^* = \sigma^2_{NXS}\left(\frac{\Delta t^*}{\Delta t_{rest}}\right)^{2\beta} = \sigma^2_{NXS}\left(\frac{\Delta t^*}{\Delta t_{obs}}\right)^{2\beta}(1+z)^{2\beta} \qquad (2)$$

Both the effects of different monitoring times and of different source redshifts are present in this expression. The proposed correction is of course especially important for high redshift sources, whose variability would otherwise be underestimated.

2. Data

To test the proposed correction, we adopt the XMM-Newton Serendipitous Source Catalogue, Data Release 3 (XMMSSC, Watson et al. 2009) already used in Vagnetti et al. 2011. In this analysis we cross correlate XMMSSC with the quasar catalogues derived by the Sloan Digital Sky Survey, Data Release 7 (Schneider et al. 2010) and Data Release 10 (Paris et al. 2014). Adopting a matching radius of 5 arcsec, we extract 871 sources with observations in at least 2 epochs, for a total of 2683 observations (sample A). The histogram of the rest-frame time intervals is shown in Fig.1. We use the XMM-Newton band EP9 (0.5-4.5 keV).

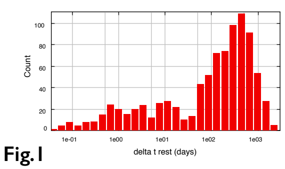

Fig.1

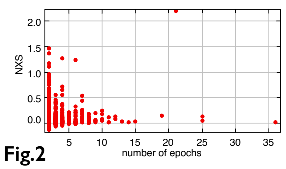

Fig.2

We show in Fig.2 the NXS values as a function of the number of epochs and notice that they converge to a nearly uniform level (with one exception for a strongly variable and flaring source) for relatively large number of epochs, while deviate a lot for small numbers, including also negative values. We have verified that a similar trend is obtained extracting N random data from a well sampled light-curve. It is thus desirable to exclude sources with a small number of observations. For the purpose of obtaining a sample sufficiently representative for our analysis, we choose a reasonable compromise, removing sources with only 2 epochs, and/or with NXS<0. The resulting sample includes 284 sources with a total of 1402 observations (sample B).

3. Results and discussion

To apply the correction of the cosmological bias to the NXS values, we have computed the structure function for sample B. This is shown in Fig. 3 and its slope is $\beta = 0.10 \pm 0.01$, equal to the value found by Vagnetti et al. 2011. In Fig. 4, we plot the NXS values as a function of the rest-frame length of the monitoring time (blue dots). There is evidence of an increasing trend with Δt, with slope 0.20±0.06, correlation coefficient r=0.21, and probability $P(>r)=4 \cdot 10^{-3}$. We correct the NXS values according to Eq. 2, choosing a fixed rest-frame time interval Δt^*=1000 days. The corrected values are plotted also in Fig. 4 as red dots; the corrected slope is about zero (0.002±0.060), with no correlation. The cosmological bias is stronger for high redshift sources, which have in average higher luminosities, therefore we expect that NXS is underestimated for high luminosity sources. In Fig. 5, we show NXS as a function of the 0.5-4.5 keV luminosity; blue dots represent uncorrected NXS values, while the red dots correspond to the corrected values at 1000 days rest-frame. Power-law fits for the two cases are shown, with slopes -0.33±0.11 and -0.31±0.11 respectively: the correction results in a slight flattening of the variability-luminosity relation. For comparison, we compute the dependence on luminosity for the variability estimates obtained by the structure function. We use the unbinned discrete structure function contributions for each pair of measurements i and j, $UDSF_{ij} = |\log f(t_i) - \log f(t_j)|$. Because some sources have much more numerous observations than others, we take average values of the UDSF for each of them, to weight the contributions of different sources uniformly. We then consider three bins of time lag ($1.5 < \log\tau \le 2$; $2 < \log\tau \le 2.5$; $2.5 < \log\tau \le 3$), computing in each of them the average UDSF value for each source. This is shown in Fig. 6, where black, blue, and red points refer to bins of increasing time lag. The fits, corresponding to a power-law trend $SF \propto L_X^{-k}$, produce k=0.23±0.11, 0.21±0.04, 0.16±0.05 for increasing time lag. To compare with the NXS result, we must consider that NXS gives average values of the variability without distinguishing the time scale. Looking at the histogram in Fig. 1, the dominant time intervals are between 300 and 1000 days. Taking the square root of NXS (usually called F_{var}), we have $F_{var} \propto L_X^{-k}$, with k=0.165±0.06, and k=0.155±0.06 for the uncorrected and corrected cases respectively. This is similar to the result of the SF for the 2.5-3 bin in log τ.

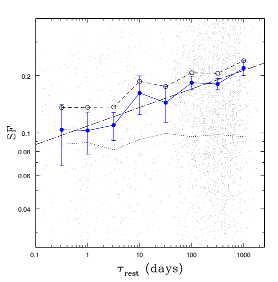

Fig.3

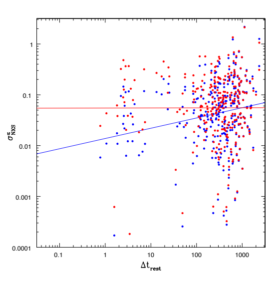

Fig.4

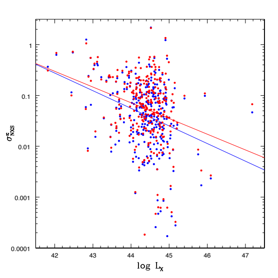

Fig.5

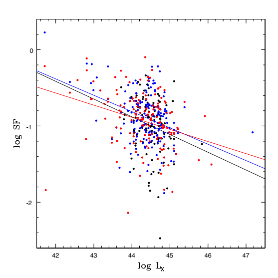

Fig.6

Bibliography

de Vries, W.H. et al. 2005, AJ 129, 615
Lanzuisi, G. et al. 2014, ApJ 781, 105
Lawrence, A. & Papadakis, I. 1993, ApJ 414, L85
Nandra, K. et al. 1997, ApJ 476, 70
Papadakis, I.E. et al. 2008, A&A 487, 475
Paris, I. et al. 2014, A&A 563, A54
Ponti, G. et al. 2012, A&A 542, A83
Schneider, D.P. et al. 2010, AJ 139, 2360
Trevese, D. et al. 1994, ApJ 433, 494
Vagnetti, F. et al. 2011, A&A 536, A84
Vaughan, S. et al. 2003, MNRAS 345, 1271
Watson, M.G. et al. 2009, A&A 493, 339